UNITED STATES DEPARTMENT OF COMMERCE
Bureau of the Census
Washington, DC 20233-0001

March 30,1999

DSSD CENSUS 2000 PROCEDURES AND OPERATIONS MEMORANDUM SERIES R-4

MEMORANDUM FOR Dennis Stoudt
Assistant Division Chief, Processing Systems
Decennial Systems and Contracts Management Office

From: Donna Kostanich DK
Assistant Division Chief, Sampling and Estimation
Decennial Statistical Studies Division

Prepared by: Randy ZuWallack RZ
Sample Design Team
Decennial Statistical Studies Division

Subject: Accuracy and Coverage Evaluation (ACE) Survey: Sample Summary File and Sample Design File Documentation

This memorandum documents the layout of two files that will be continuously updated during the sample selection for the Census 2000 Accuracy and Coverage Evaluation (ACE). The first is the ACE Sample Summary File, which will contain cluster and housing unit totals at the state level. This file will aid in the monitoring of the sampling procedures by providing expected and actual results which will then be compared to identify extreme differences. Attachment A contains a file layout for the Sample Summary File. The second file is the ACE Sample Design File. This file tracks the path that each block cluster travels during the ACE sampling procedures. The Sample Design File contains categorical variables corresponding to each procedure as well as parameters and housing unit totals. In addition, sampling weights will be assigned based on the final path each cluster follows during the ACE sampling operations. Attachment B contains a file layout for the Sample Design File. Together the Sample Summary File and Sample Design File will document the history of the ACE design and serve as a reference during evaluations and estimation.

The creation of the Sample Summary File will occur following the creation of the Universe File[1]. The Sample Design File will be created following the block cluster sampling[2]. The Sample Summary File and Sample Design File will be updated in the specifications for each of the ACE sampling procedures, which include the initial ACE block cluster sampling, the ACE block cluster reduction, small block subsampling, large block subsampling, and E-sample identification. Although the Sample Summary File and Sample Design File will be updated following each of these processes, the layout for these files will be documented in this specification. A source code is assigned to each variable indicating where in the processing the variable is first encountered. These source codes are listed following each file layout. For information not foreseen as being required for the sampling procedures, space will be left for additions to the files. This space will be filled as necessary following each process, and will be documented in the specification for that process. At the conclusion of all ACE sampling operations, the final layout for the Sample Summary File and Sample Design File will be documented.

For questions concerning the Sample Design File or the Sample Summary File, contact Deborah Fenstermaker 301-457-4195 or Randy ZuWallack 301-457-1963.

cc: DSSD Census 2000 Procedures and Operations Memorandum Series Distribution List
ACE Implementation Team/Statistical Design Team Leaders List
Sample Design Team

---

[1]Memorandum from Kostanich to Stoudt, "Accuracy and Coverage Evaluation Survey: Universe File and Sampling Parameter File Specification", March 1999.

[2]Memorandum from Kostanich to Stoudt, "Accuracy and Coverage Evaluation Survey: Block Cluster Sample Selection", March 1999.

Attachment A

Sample Summary File

The Sample Summary File contains one record for each of the 50 states, the District of Columbia and Puerto Rico for a total of 52 records. The initial version of the file, which will be created following the creation of the Universe File, is called ACE2000_SSFV1.<mmddyy>. The extension <mmddyy> is the date the file is created (i.e. 123199 is the extension for a file created on December 31, 1999). For each subsequent update to the file, the version number will increase by one (i.e. ACE2000_SSFV2.<mmddyy>, ACE2000_SSFV3.<mmddyy>). The layout for the Sample Summary File is as follows:

Variable Description	Name	Places	Source
Census Region	REGION	1	UN
Census Division	DIV	2	UN
State code (01-72 = FIPS State Code)	STATE	3-4	UN
Number of HUs budgeted for listing in med. and lg. clusters	BLIST	6-13	UN
Target number of clusters in small sampling strata	TCLUSTS	15-17	UN
Target number of clusters in med. and lg. sampling strata	TCLUST	19-22	UN
Target number of clusters in AIR sampling strata	TCLUSTA	24-26	UN
Total number of block clusters	NCLUST	28-35	BC
Total number of HUs	NHU	37-44	BC
Expected clusters in sample to list	ECLUSTL	46-49	UN
Expected HUs in sample to list	EXPHUL	51-58	UN
Additional space		59-80	
Clusters in sample to list after 1st step sampling	NCLUSTL1	81-85	CS
Estimated HUs in sample to list after 1st step sampling	NHUL1	87-94	CS
Estimated HUs in sample to list after 1st step sampling in Med & Lg clusters	NHUL1_ML	96-103	
Indicator for second step of block cluster sampling 1 = Second step needed, 2 = Second step not needed	I2	105	CS
Clusters in sample to list after 2nd step sampling	NCLUSTL2	107-111	CS
Estimated HUs in sample to list after 2nd step sampling	NHUL2	113-120	CS
Estimated HUs in sample to list after 2nd step sampling in Med & Lg clusters	NHUL2_ML	122-129	CS
Additional space		130-150	
Preliminary Number of HUs on Independent List	NHUILLP	151-158	AR
Number of Housing Units On the DMAF	NHUDMAF	160-167	AR
Additional space reserved for ACE reduction		168-270	
Number of HUs on Independent List	NHUILL	271-278	SB
Expected number of clusters selected for ACE	ECLUST	280-284	SB
Expected number of Independent List HUs for ACE	EHUIL	286-293	SB
Number of clusters selected for ACE	NCLUST	295-299	SB
Number of Independent List HUs for ACE	NHUIL	301-308	SB
Additional space		309-330	

Attachment A

Variable Description	Name	Places	Source
Number of HUs on the Preliminary Enhanced List	NHUEL	331-338	LB
Number of ACE HUs on the Preliminary Enhanced List	NHUELA	340-347	LB
Number of non-ACE HUs on the Preliminary Enhanced List	NHUELN	349-346	LB
Expected number of HUs for interview	EHUINT	358-365	LB
Expected number of ACE HUs for interview	EHUINTA	367-374	LB
Expected number of non-ACE HUs for interview	EHUINTN	376-383	LB
Number of HUs for interview	NHUINT	385-392	LB
Number of ACE HUs for interview	NHUINTA	394-401	LB
Number of non-ACE HUs for interview	NHUINTN	403-410	LB
Additional space		411-430	
Number of CUF HUs	NHUCUF	431-438	ES
Number of CUF HUs in block cluster with an ESPS code of 1	NHUCUF1	440-447	ES
Number of CUF HUs in block cluster with an ESPS code of 2	NHUCUF2	449-456	ES
Expected number of E-sample HUs	EHUES	458-465	ES
Expected number of E-sample HUs with an ESPS code of 1	EHUES1	467-474	ES
Expected number of E-sample HUs with an ESPS code of 2	EHUES2	470-483	ES
Number of E-sample HUs	NHUES	485-492	ES
Number of E-sample HUs with an ESPS code of 1	NHUES1	494-501	ES
Number of E-sample HUs with an ESPS code of 2	NHUES2	503-510	ES
Additional Space		511-600	

Source Codes

AR: ACE Reduction
BC: Block Clustering
CS: Block Cluster Sampling
ES: E-sample Identification
LB: Large Block Subsampling
SB: Small Block Subsampling
UN: Universe File Creation

Attachment B

Sample Design File

The Sample Design File contains one record per block cluster selected during the initial block cluster sampling. If the block clusters falls out of sample during the second step of sampling or during small block subsampling, the remaining variables will be left blank. The initial version of the file, which will be created following the initial block cluster selection, is called ACE2000_SDFV1.<mmddyy>. For each subsequent update to the file, the version number will increase by one (i.e. ACE2000_SDFV2.<mmddyy>, ACE2000_SDFV3.<mmddyy>). The layout for the Sample Design File is as follows:

Variable Description	Name	Places	Source
Census Region	REGION	1	UN
Census Division	DIV	2	UN
State code	STATE	3-4	UN
County code	COUNTY	5-7	UN
Local census office	LCO	8-11	CS
Interim Tract (Pseudo Tract)	ITRACT	12-17	BC
Current Sample Indicator	CSI	19	UO
ACE block cluster number	CLUST	21-25	CS
Check Digit	DIGIT	26	CS
Geography block cluster number	GCLUST	28-32	BC
Type of Enumeration Area Recode	TEACR	34	CS
Type of Enumeration Area group	TEAG	36	BC
Number of HUs used for sample design	NHU	37-41	BC
Number of MAF HUs	NHUM	43-47	BC
Number of 1990 HUs	NHU90	49-53	BC
Sampling Stratum 1 = Small 2 = Medium 3 = Large 4 = American Indian Reservation	SS	55	UN
American Indian Country Indicator 0 = No American Indian Country 1 = American Indian Reservation/trust land 2 = Tribal Jurisdiction Area/ Alaska Native Village Statistical Area/ Tribal Designated Statistical Area	AICIND	56	BC
Demographic/Tenure Group code	DTCODE	57-58	UN
Demographic/Tenure Group label	DTLABEL	59-60	UN
Estimated Urbanicity of block cluster 1 = Urban Area with population ≥250,000 2 = Other Urban Area 3 = Non-Urban Area	ECLUSURB	62	UN

Attachment B

Variable Description	Name	Places	Source
Size Category 1=Small (0-2 hus) 2=Medium (3-79 hus) 3=Large (80+ hus)	SIZCAT	63	UN
Additional space		64-91	
First step index number	INDEX1	92-99	CS
Initial block cluster sampling Indicator 1 = Selected	BC1	101	CS
Random Start for initial block cluster sampling	RS1	103-113	UN
Take-every for initial block cluster sampling	TE1	115-125	UN
Second block cluster sampling Indicator 0 = Not Selected, 1 = Selected	BC2	127	CS
Random Start for second block cluster sampling	RS2	129-139	CS
Take-every for second block cluster sampling	TE2	141-151	CS
Unbiased weight after block cluster sampling	WEIGHTBC	153-164	CS
Additional space		165-175	
Preliminary Number of HUs on the Independent List	NHUILP	176-180	AR
Number of Housing Units On the DMAF	NHUDMAF	182-186	AR
Additional space reserved for ACE reduction		187-277	
Unbiased weight after ACE reduction	WEIGHTAR	278-289	AR
Additional space		290-300	
Number of HUs on the Independent List	NHUIL	301-305	SB
Independent List Cluster Category	ILCC	307	SB
Small Block Subsampling Indicator 0 = Not Selected, 1 = Selected	SB	308	SB
Random Start for Small Block subsampling	RSSB	310-320	SB
Take-every for Small Block subsampling	TESB	322-332	SB
Unbiased weight for ACE cluster	WEIGHTC	334-345	SB
Additional space		346-370	
Relisted Block Cluster Flag 0 = Not Relisted, 1 = Relisted	RELIST	371	LB
Number of total hus on the EL in block cluster	NHUEL	373-377	LB
Number of ACE hus on the EL in cluster	NHUELA	379-383	LB
Number of non-ACE hus on the EL in cluster	NHUELN	385-389	LB
Enhanced List Cluster Category 1 = NHUELI< 80 hus, 2 = NHUELI ≥ 80 hus	ELCC	391	LB
Random Start for Large Block subsampling	RSLB	393-403	LB
Take-every for Large Block subsampling	TELB	405-415	LB
Number of Segments per block cluster	NSEG	417-418	LB
Number of selected segments	NSEGSAM	420-421	LB
Day of Arrival	DAY	423-424	LB
Daily Cluster Order Number	DCON	426-429	LB

Attachment B

Variable Description	Name	Places	Source
Final Cluster Order Number	CON	431-434	LB
Non-ACE Subsampling Flag	NISUB	436	LB
Number of total hus for interview in block cluster	NINT	438-442	LB
Number of ACE hus for interview in block cluster	NINTA	444-448	LB
Number of non-ACE HUs for interview	NINTN	450-454	LB
Unbiased weight for P-sample HUs	WEIGHTP	456-467	LB
Additional space		468-490	
Number of CUF HUs in block cluster with an ESPS code of 1	NHUCUF1	491-495	ES
Number of CUF HUs in block cluster with an ESPS code of 2	NHUCUF2	497-501	ES
Number of CUF HUs in block cluster	NHUCUF	503-507	ES
Number of CUF HUs in selected segments with an ESPS code of 1	NHUCUFS1	509-513	ES
Number of CUF HUs in selected segments with an ESPS code of 2	NHUCUFS2	515-519	ES
Number of CUF HUs in selected segments of a block cluster	NHUCUFS	521-525	ES
E-Sample Identification cluster category 1 = NHUCUF < 80 2 = NHUCUF ≥ 80 and NHUCUFS < 80 3 = NHUCUF ≥ 80 and NHUCUFS ≥ 80 4 = NHUCUF ≥ 80 and RELIST = 1 5 = NHUCUF ≥ 80 and List/Enumerate	EICC	527	ES
Random Start for E-sample subsampling	RSES	529-539	ES
Take-every for E-sample subsampling	TEES	541-551	ES
Number of E-sample HUs in block cluster with an ESPS code of 1	NHUES1	553-557	ES
Number of E-sample HUs in block cluster with an ESPS code of 2	NHUES2	559-563	ES
Number of E-sample HUs in block cluster	NHUES	565-569	ES
Unbiased weight for E-sample HUs with an ESPS code of 1	WEIGHTE1	571-582	ES
Unbiased weight for E-sample HUs with an ESPS code of 2	WEIGHTE2	584-595	ES
Additional Space		596-620	
Trimmed weight for P-sample HUs	TRIMWTP	621-632	WT
Trimmed weight for E-sample HUs with an ESPS code of 1	TRIMWTE1	634-645	WT
Trimmed weight for E-sample HUs with an ESPS code of 2	TRIMWTE2	647-658	WT
Additional Space		659-750	

Source Codes

AR: ACE Reduction
BC: Block Clustering
CS: Block Cluster Sampling
ES: E-sample Identification
LB: Large Block Subsampling
SB: Small Block Subsampling
UN: Universe File Creation
UO: Updated for each operation
WT: Weight Assignment